

Importing Data into SAS On Demand for Academics (SODA)

Using SAS Studio



Dataset

- This tutorial is a walkthrough with a sample set of data. You may use this to walk through the tutorial, if you wish, but for your assignments, you will be asked to use your own dataset (as specified within the course).

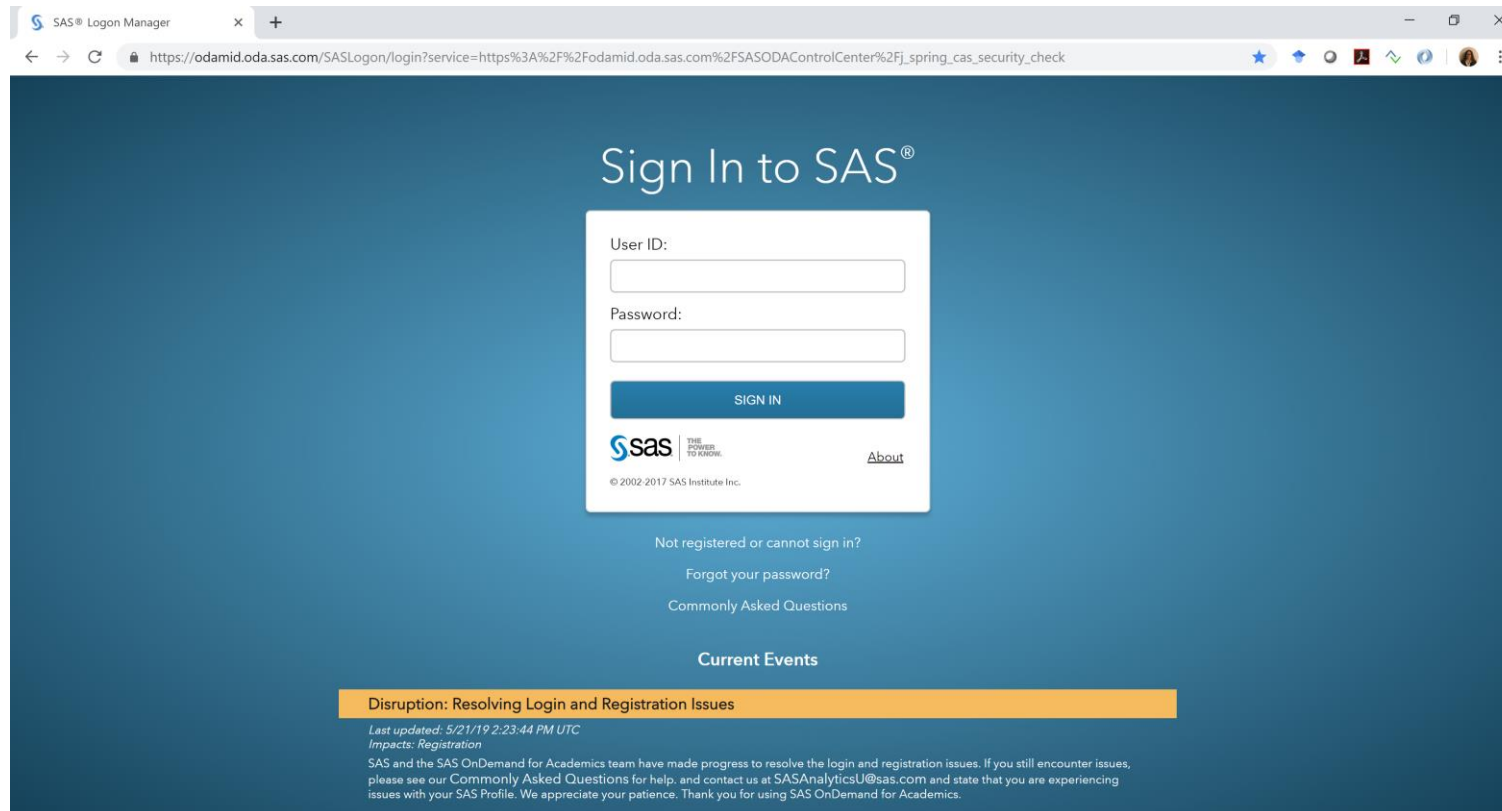
Dataset reference:

Skoryk, M. (2021). Sepsis Prediction from Clinical Data. Version 1.
Retrieved from <https://www.kaggle.com/maxskoryk/datasepsis>



Access the SAS OnDemand for Academics Control Center

<https://odamid.oda.sas.com/SASODAControlCenter>



The screenshot shows a web browser window with the title "SAS® Logon Manager". The address bar displays the URL: https://odamid.oda.sas.com/SASLogon/login?service=https%3A%2F%2Fodamid.oda.sas.com%2FSASODAControlCenter%2Fj_spring_cas_security_check. The main content area has a dark blue background with the text "Sign In to SAS®" in white. Below this, there is a white login form with two input fields: "User ID:" and "Password:". A blue "SIGN IN" button is positioned below the password field. At the bottom of the form, the SAS logo is displayed with the tagline "THE POWER TO KNOW." and a link to "About". Below the login form, there are links for "Not registered or cannot sign in?", "Forgot your password?", and "Commonly Asked Questions". At the bottom of the page, there is a section titled "Current Events" with a yellow banner that reads "Disruption: Resolving Login and Registration Issues". Below the banner, it states "Last updated: 5/21/19 2:23:44 PM UTC" and "Impacts: Registration". A paragraph of text follows, stating that SAS and the SAS OnDemand for Academics team have made progress to resolve the login and registration issues, and provides contact information for further assistance.

Sign In to SAS®

User ID:

Password:

SIGN IN

SAS THE POWER TO KNOW. [About](#)

© 2002-2017 SAS Institute Inc.

Not registered or cannot sign in?
Forgot your password?
Commonly Asked Questions

Current Events

Disruption: Resolving Login and Registration Issues

Last updated: 5/21/19 2:23:44 PM UTC
Impacts: Registration

SAS and the SAS OnDemand for Academics team have made progress to resolve the login and registration issues. If you still encounter issues, please see our Commonly Asked Questions for help, and contact us at SASAnalyticsU@sas.com and state that you are experiencing issues with your SAS Profile. We appreciate your patience. Thank you for using SAS OnDemand for Academics.



SAS OnDemand for Academics (SODA) Control Center

The screenshot shows the SAS OnDemand for Academics (SODA) Control Center dashboard. The browser address bar displays `https://odamid.oda.sas.com/SASODAControlCenter/`. The page header includes the SAS logo and the text "SAS OnDemand for Academics Control Center - United States". A blue navigation bar at the top contains the word "Dashboard" and a user profile for "Stefanie Reay".

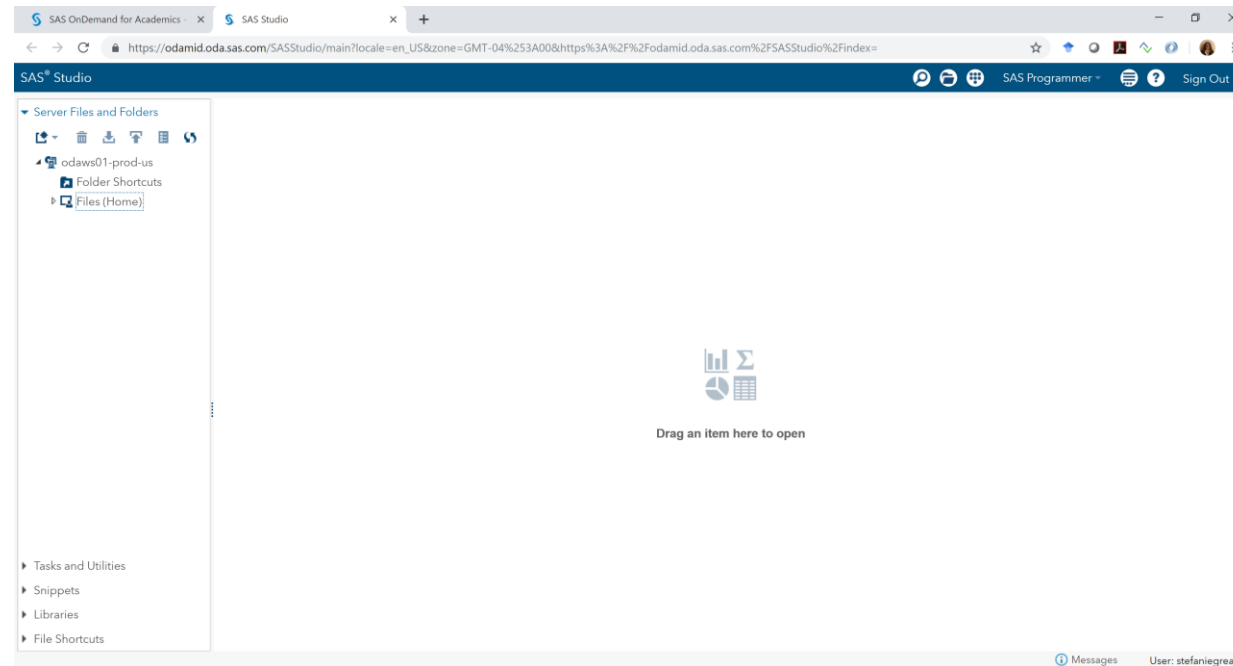
The main content area is divided into several sections:

- Current Events** (click entry for details):
 - Resolving Login and Registration Issues**: Last updated: May 21, 2019 14:23 UTC. Impacts: Registration, Login Page.
- Planned Events** (click entry for details):
 - Scheduled Maintenance on Tuesday, July 16, 8:30 - 10:00 am EDT**: Last updated: Jul 09, 2019 17:38 UTC. Scheduled start time: Jul 16, 2019 12:30 UTC. Scheduled end time: Jul 16, 2019 14:00 UTC.
- Applications** (tabbed view):
 - SAS Studio**: Write and run SAS code with a Web-based SAS development environment. Actions: Clear my saved tabs.
 - SAS Enterprise Guide**: Deliver the power of SAS from an easy-to-use, point-and-click interface. (Download Required)
 - JMP access to SAS Servers (U.S. users only)**: Statistical discovery software. Users must have a copy of JMP software. (Configuration Steps Required)
 - SAS Enterprise Miner**: Reveal valuable insights with powerful data mining software. (Configuration Steps Required)
- Reference**:
 - Support Site
 - Step-by-Step Registration Guides
 - User's Guide
 - Commonly Asked Questions
 - Status Page
- Quotas** (learn more):
 - Home Directory (4.4MB/5120MB): 0%
 - Course Directory (207.0MB/3072MB): 7%
- Notices**:
 - German, Spanish and French Locale Users Impacted by a SAS Studio Issue**: Last updated: Jun 18, 2019 15:02 UTC. Importance: High. Impacts: SAS Studio.

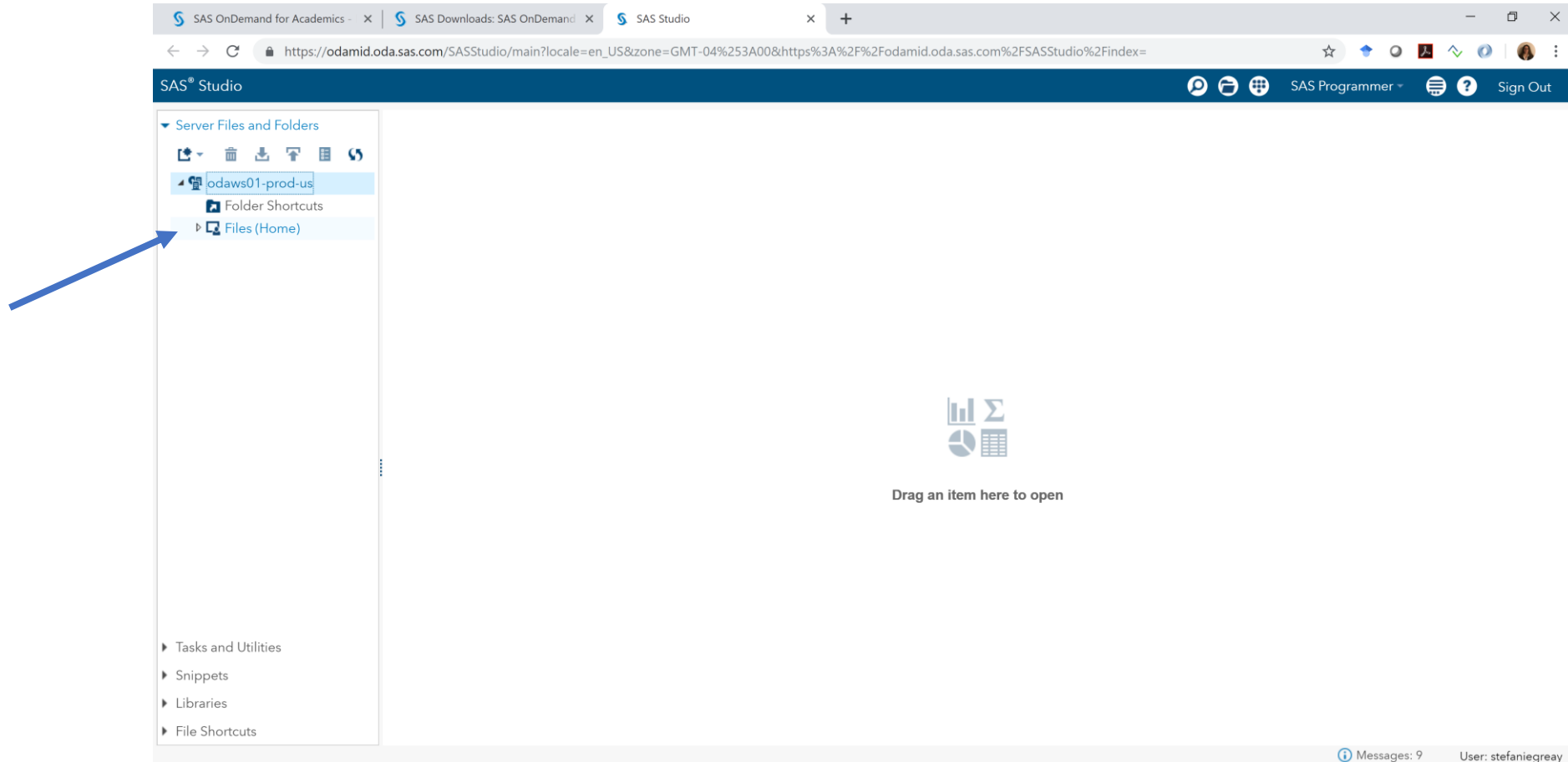


SAS Studio

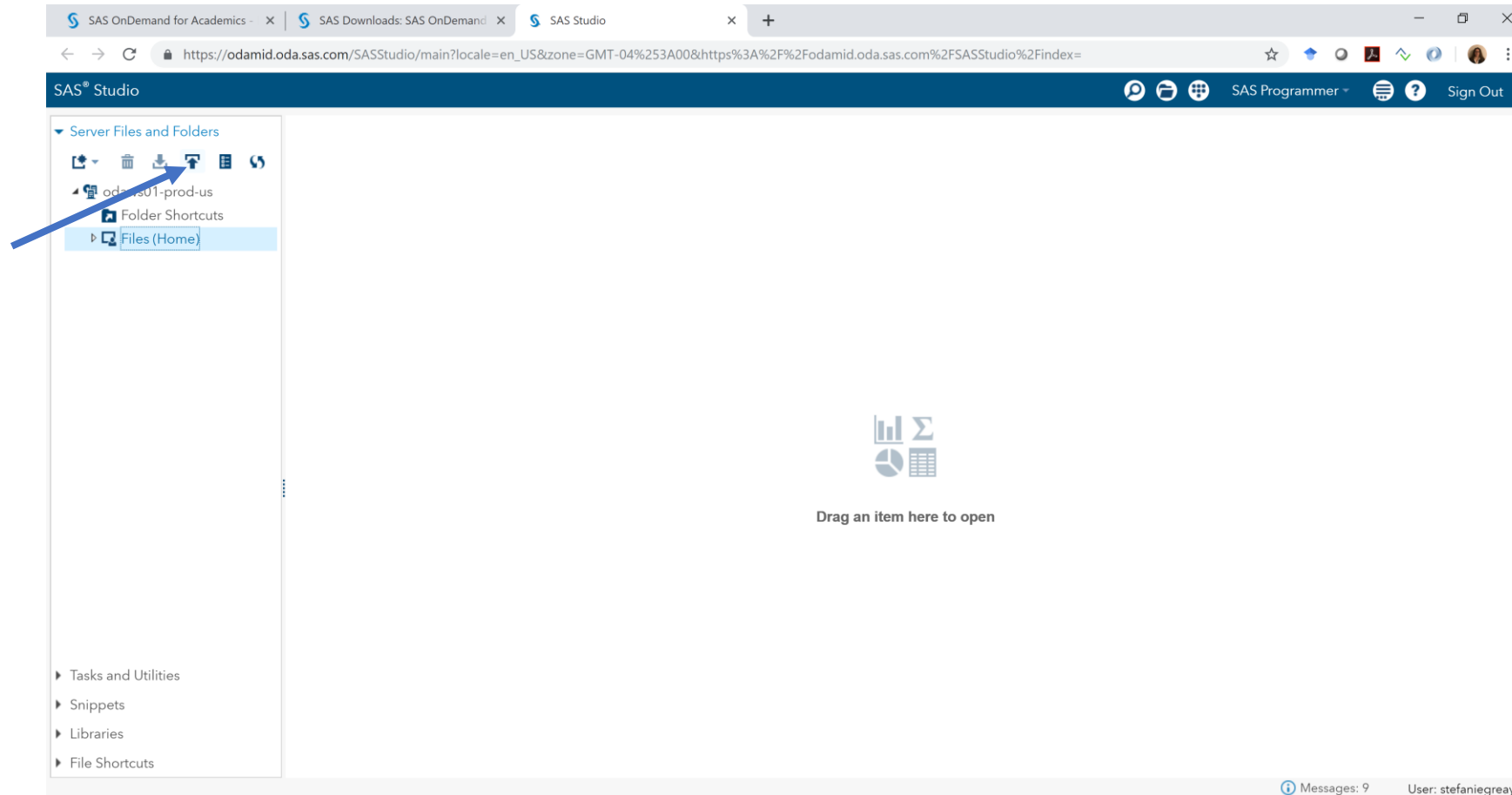
<https://odamid.oda.sas.com/SASStudio/>



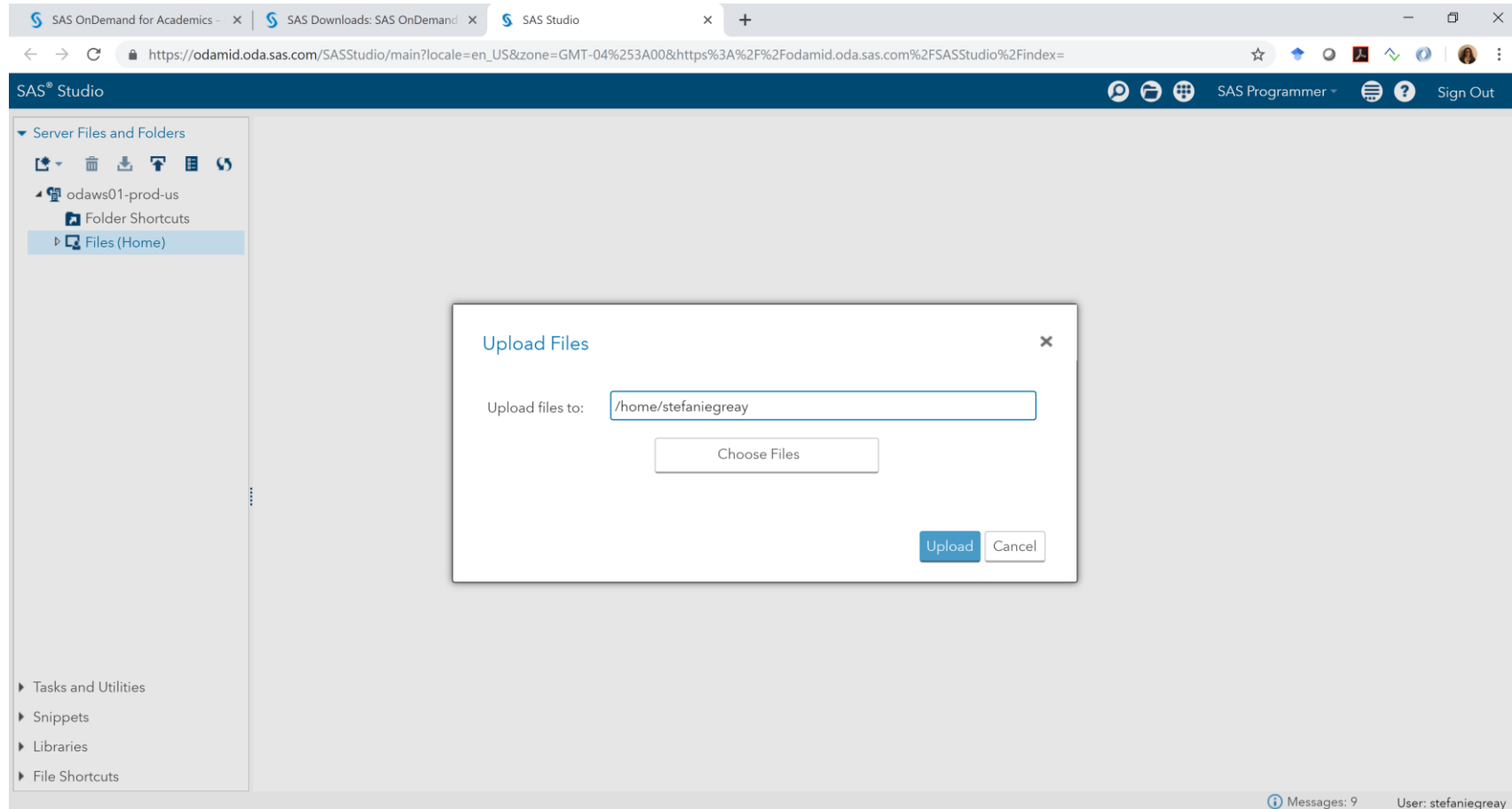
Click on Files(Home)



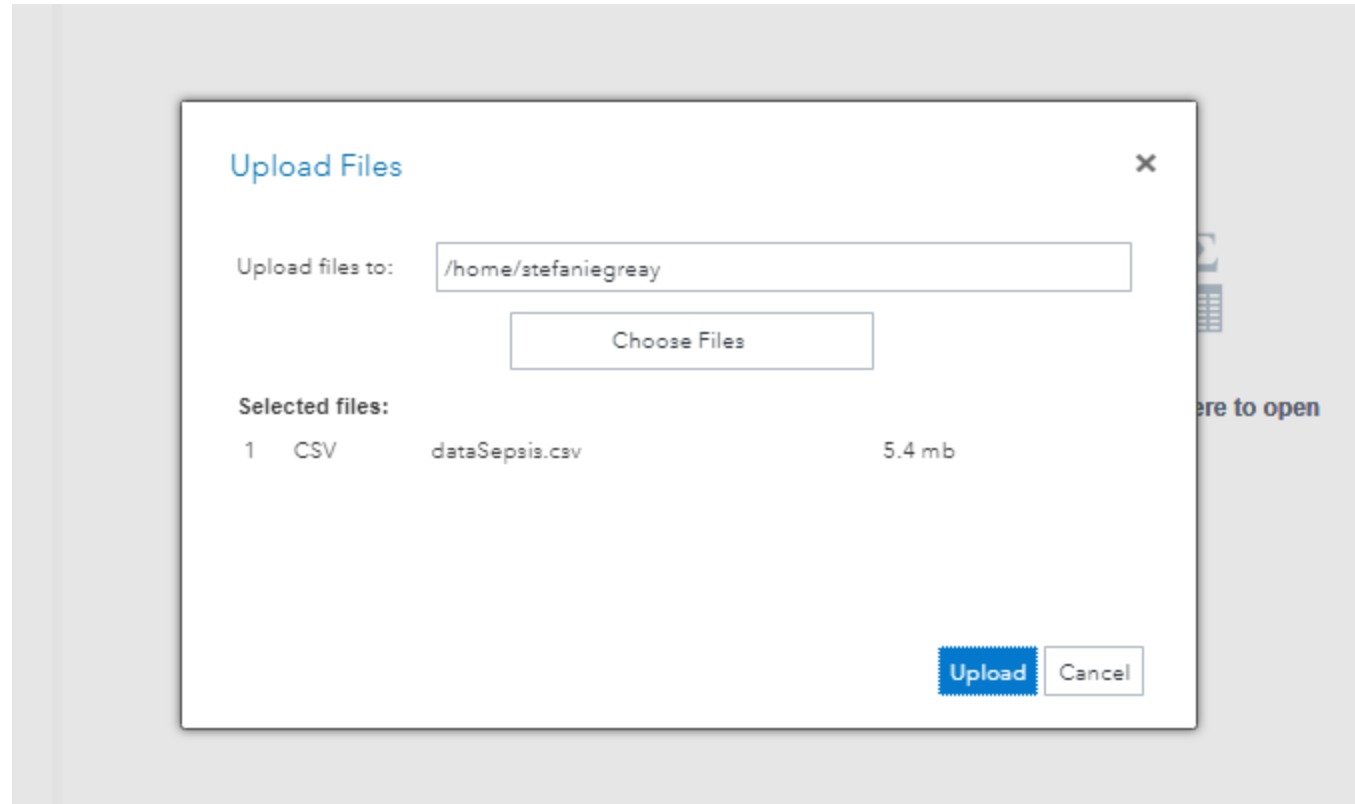
The Upload button will display in dark blue



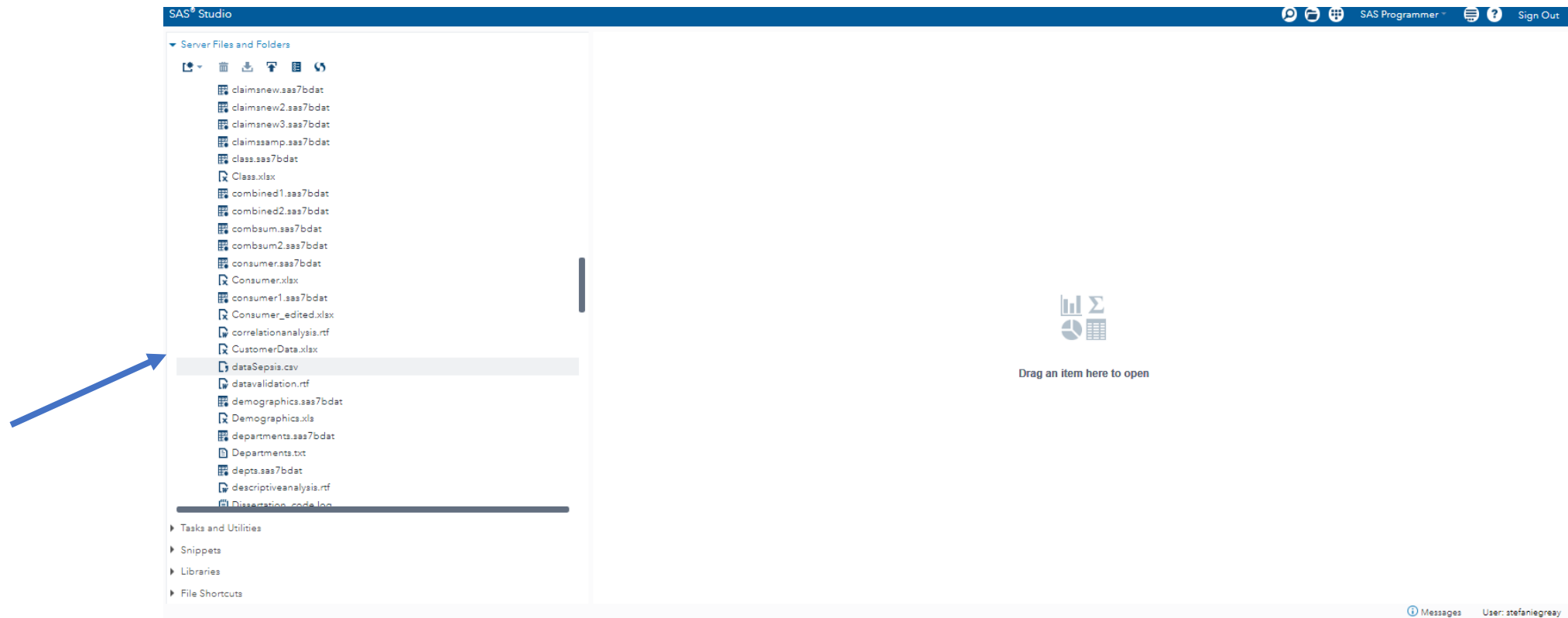
You can create a folder at this point, if you wish, or simply upload to your home directory.



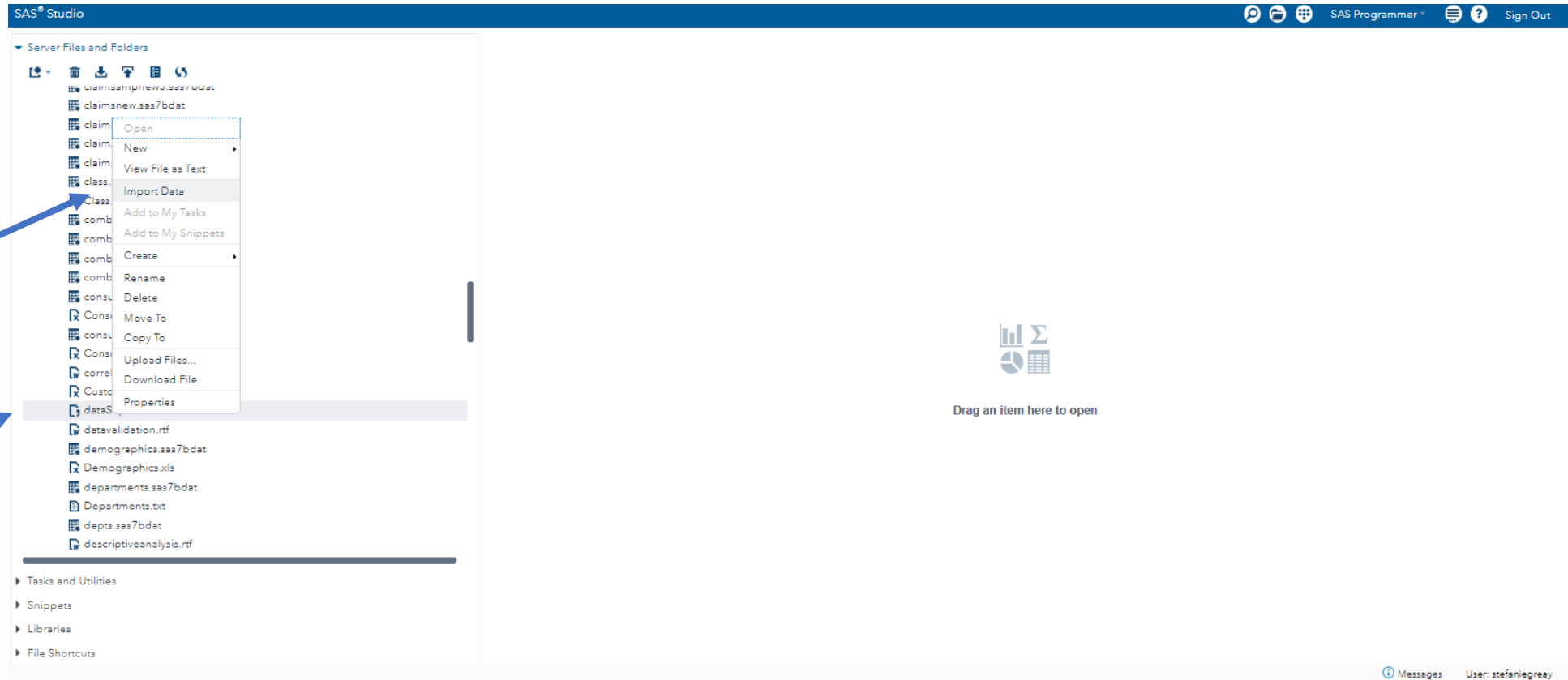
Select “Choose Files” to browse your computer for the dataset you want to upload. Once the dataset has been selected, click “Upload.”



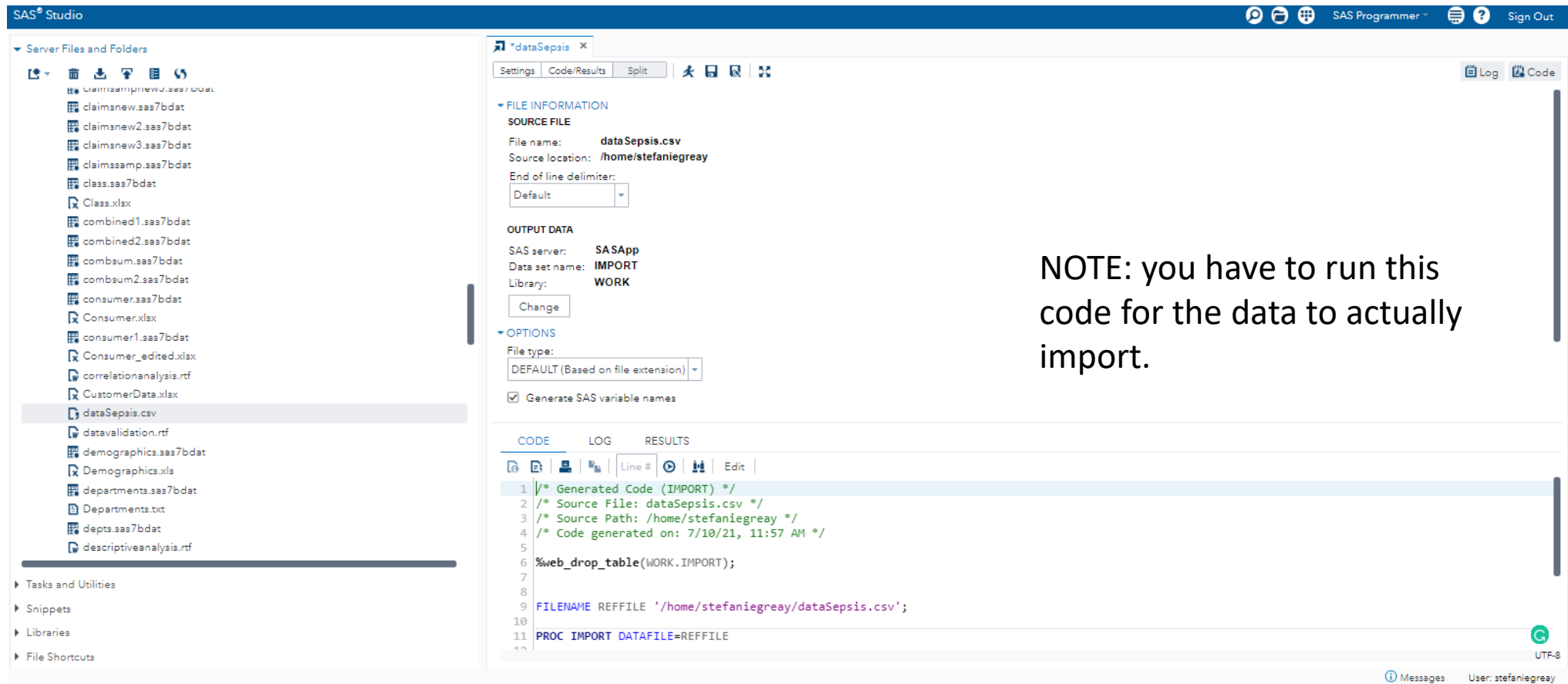
You will be able to view your files by clicking on “Files(Home)” to verify that your file successfully uploaded.



To import the dataset into a SAS dataset format (from the current csv format), right click on the name of the file, and select “Import Data.”



The Proc Import code will be written for you (save this as a template to use for future imports!)



NOTE: you have to run this code for the data to actually import.



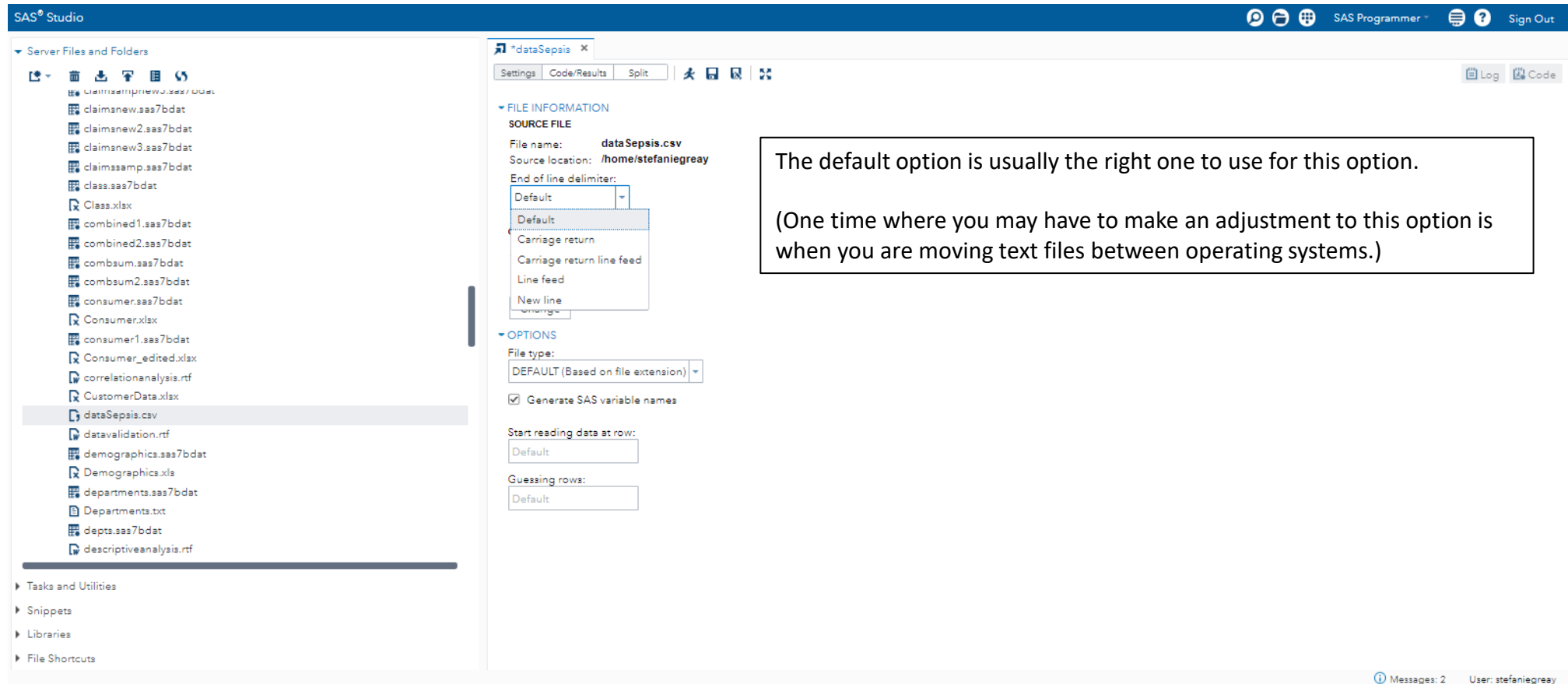
Default Import Options

The screenshot shows the SAS Studio interface. On the left, the 'Server Files and Folders' pane lists various files, with 'dataSepsis.csv' highlighted. On the right, the 'Import Data' dialog is open, showing the 'FILE INFORMATION' and 'OPTIONS' tabs. Blue arrows point from numbered list items to specific settings in the dialog:

- 1. End of line delimiter - what character is used to identify the end of each line or row of data? (Points to the 'End of line delimiter' dropdown menu.)
- 2. Output Data - where do you want to store your data? (Work is a temporary space, and import is the default name of the dataset.) (Points to the 'Data set name' dropdown menu.)
- 3. File Type - what type of file are you importing (the default setting will automatically select the file type based on the file extension, like .csv, .xlsx, or .txt, for example) (Points to the 'File type' dropdown menu.)
- 4. Generate SAS variable names - do you want column names generated from the first row of data? (Points to the 'Generate SAS variable names' checkbox.)
- 5. Start reading data at row - what row do you want to be your first row in your dataset (depending on whether option 4 is checked or not) (Points to the 'Start reading data at row' input field.)
- 6. Guessing rows - what row do you want SAS to use to determine the formats, lengths, etc. of each variable or column? (Points to the 'Guessing rows' input field.)



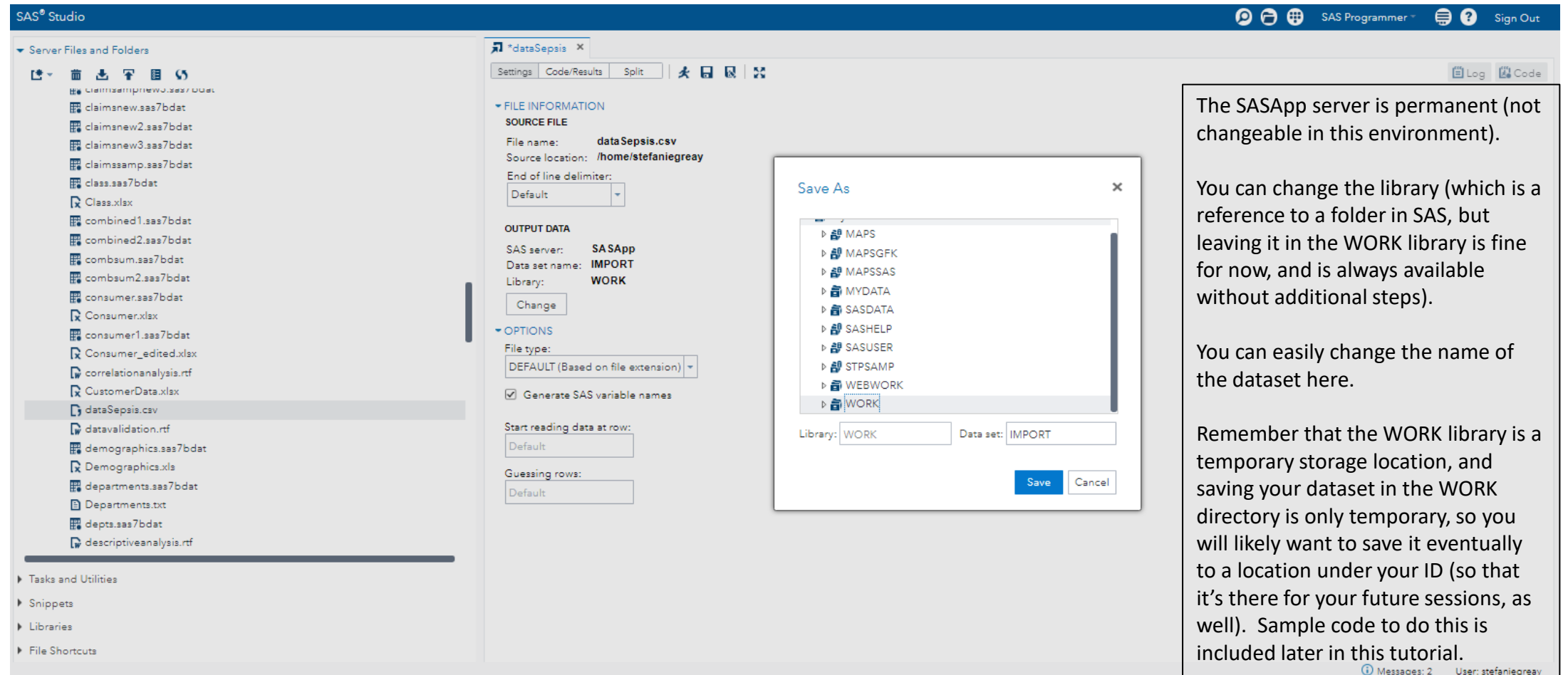
Alternative Import Options: End of Line Delimiter



The screenshot shows the SAS Studio interface. On the left, the 'Server Files and Folders' pane lists various files, with 'dataSepsis.csv' highlighted. On the right, the 'Options' pane for 'dataSepsis.csv' is displayed. The 'End of line delimiter' dropdown menu is open, showing options: 'Default', 'Carriage return', 'Carriage return line feed', 'Line feed', and 'New line'. A text box on the right side of the interface states: 'The default option is usually the right one to use for this option. (One time where you may have to make an adjustment to this option is when you are moving text files between operating systems.)' The 'File type' is set to 'DEFAULT (Based on file extension)' and 'Generate SAS variable names' is checked. The 'Start reading data at row' and 'Guessing rows' are both set to 'Default'.



Alternative Import Options: Output Data



The SASApp server is permanent (not changeable in this environment).

You can change the library (which is a reference to a folder in SAS, but leaving it in the WORK library is fine for now, and is always available without additional steps).

You can easily change the name of the dataset here.

Remember that the WORK library is a temporary storage location, and saving your dataset in the WORK directory is only temporary, so you will likely want to save it eventually to a location under your ID (so that it's there for your future sessions, as well). Sample code to do this is included later in this tutorial.



Alternative Import Options: File Type

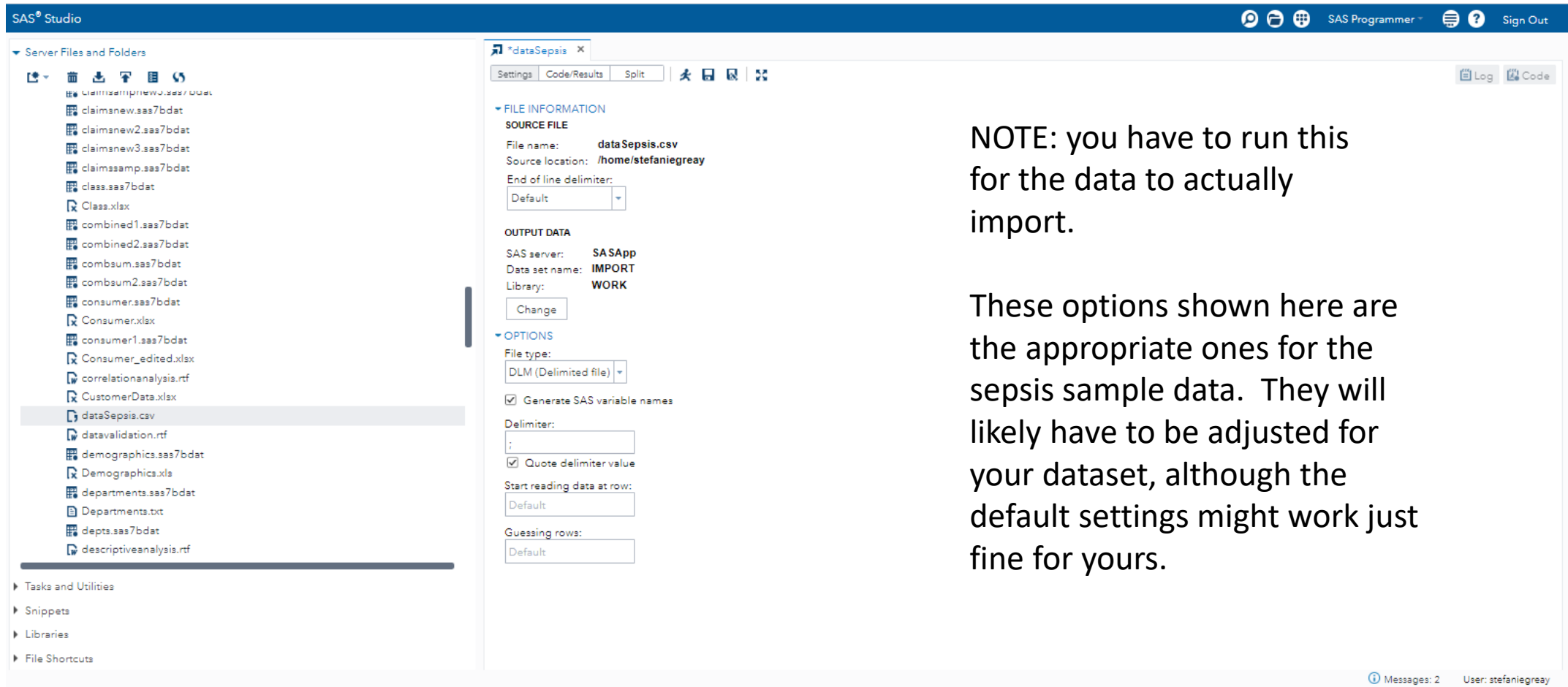
The file type option allows you to specify the specific type of file you are attempting to import, and is useful if SAS is not successful in determining it and importing the data using the DEFAULT option.

The most commonly used file types used include:

- 1) CSV (which SAS assumes is delimited or separated by commas – between the fields/variables)
- 2) DLM (a general option for any file that is delimited, or uses one or more symbols or characters to separate fields/variables (this could be a comma or could be a pipe | or a semicolon ; or other character). (You would most likely choose this option if the extension on your file was .txt or csv, if the delimiter was not a comma.)
- 3) EXCEL, XLS, and XLSX (all formats of Microsoft Excel files...depending on the format and the version of Excel that the dataset was created in or saved from, you can try different ones of these to find the correct option)



The Proc Import code will be written for you (save this as a template to use for future imports!)



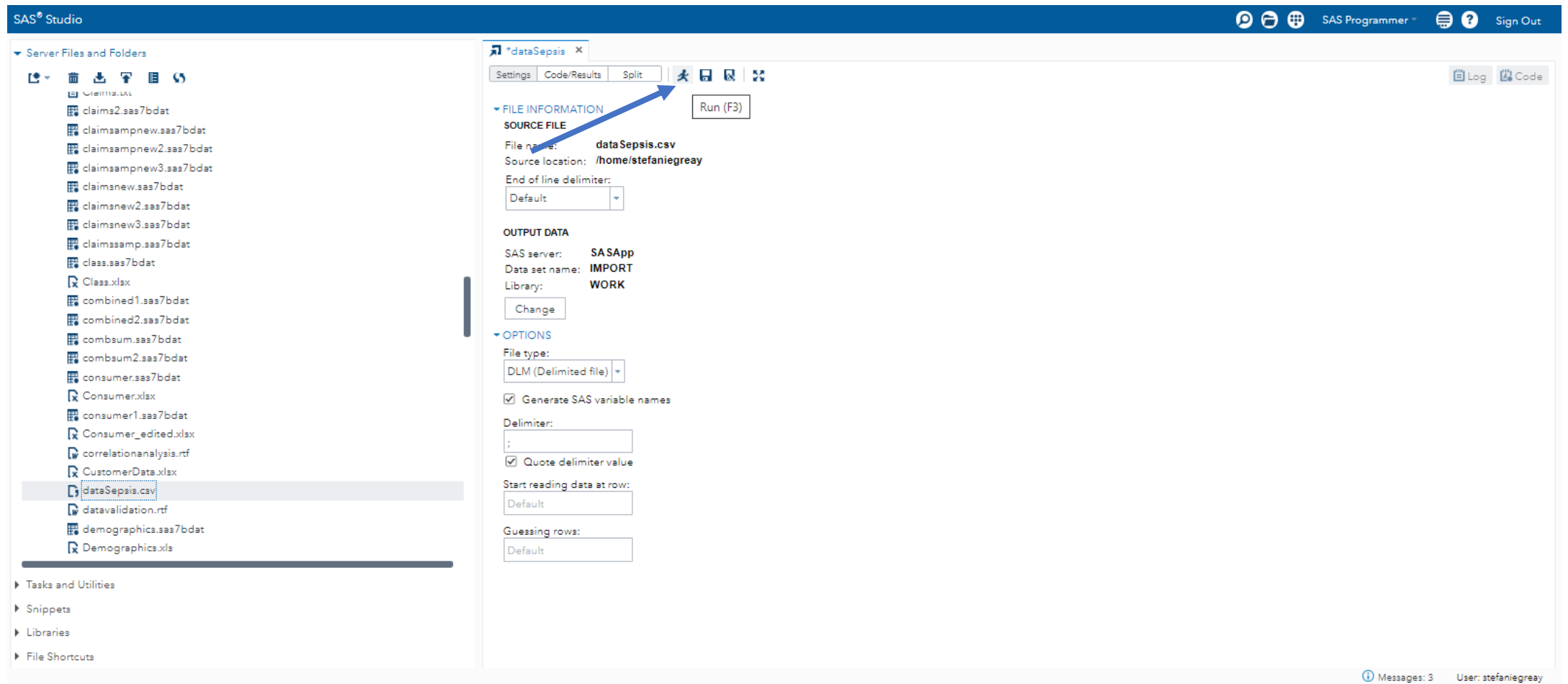
The screenshot shows the SAS Studio interface. On the left, the 'Server Files and Folders' pane lists various files, with 'dataSepsis.csv' selected. The main window displays the 'Proc Import' wizard for 'dataSepsis.csv'. The 'FILE INFORMATION' section shows the source file name and location. The 'OUTPUT DATA' section shows the SAS server, data set name, and library. The 'OPTIONS' section shows the file type as 'DLM (Delimited file)', with checkboxes for 'Generate SAS variable names' and 'Quote delimiter value' checked. The 'Delimiter' is set to ';', and the 'Start reading data at row' and 'Guessing rows' are both set to 'Default'.

NOTE: you have to run this for the data to actually import.

These options shown here are the appropriate ones for the sepsis sample data. They will likely have to be adjusted for your dataset, although the default settings might work just fine for yours.



To run the code, click the icon that looks like a guy running.



When you run the import, you will see the dataset and summary in the output data window when you click the “Code/Results” or “Split” tab and then “Output Data” and can verify its success.

SAS® Studio

Server Files and Folders

- class.sas7bdat
- Class.xlsx
- combined1.sas7bdat
- combined2.sas7bdat
- combsum.sas7bdat
- combsum2.sas7bdat
- consumer.sas7bdat
- Consumer.xlsx
- consumer1.sas7bdat
- Consumer_edited.xlsx
- correlationanalysis.rtf
- CustomerData.xlsx
- dataSepsis.csv**
- datavalidation.rtf
- demographics.sas7bdat
- Demographics.xls
- departments.sas7bdat
- Departments.txt
- depts.sas7bdat
- descriptiveanalysis.rtf
- Dissertation_code.log
- Dissertation_code.sas
- Dissertation_code_01062020_243pm.log
- Dissertation_code_01062020_243pm.sas
- Dissertation_code_01082020_1121am.sas

Tasks and Utilities

- Snippets
- Libraries
- File Shortcuts

*dataSepsis

Settings | Code/Results | Split | Log | Code

CODE | LOG | RESULTS | **OUTPUT DATA**

Table: WORK.IMPORT | View: Column names | Filter: (none)

Columns

- Select all
- ☒ HR
- ☒ O2Sat
- ☒ Temp
- ☒ SBP
- ☒ MAP
- ☒ DBP
- ☒ Resp
- ☒ EtCO2
- ☒ BaseExcess
- ☒ HCO3
- ☒ FiO2
- ☒ pH

Property | Value

Label

Name

Length

Type

Format

Inform

Total rows: 36302 Total columns: 41

	HR	O2Sat	Temp	SBP	MAP	DBP	Resp	EtCO2	BaseExcess	HCO3	FiO2	pH	PaCO2	SaO2	AST	BUN	Alkalinephos	Cr
1	103	90	NaN	NaN	NaN	NaN	30	NaN	21	45	NaN	7.37	90	91	16	14	98	9.
2	58	95	36.11	143	77	47	11	NaN	NaN	22	NaN	NaN	NaN	NaN	NaN	100	NaN	7.
3	91	94	38.5	133	74	48	34	NaN	NaN	31	0.8	NaN	NaN	NaN	NaN	30	NaN	10.
4	92	100	NaN	NaN	NaN	NaN	NaN	NaN	NaN	29	NaN	NaN	NaN	NaN	NaN	9	NaN	Na
5	155.5	94.5	NaN	147.5	102	NaN	33	NaN	-12	13	1	7.22	36	NaN	452	68	88	5.
6	73	99	36.06	100	67	49.5	16.5	NaN	-8	16	NaN	7.27	37	NaN	NaN	28	NaN	7.
7	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	0	25	NaN	7.35	48	NaN	NaN	NaN	NaN	Ni
8	82	100	35.5	112	79.5	63	14	NaN	0	23	1	7.42	37	NaN	NaN	18	NaN	Ni
9	89	100	NaN	141	85	57	17	NaN	1	25	NaN	7.43	37	NaN	NaN	9	NaN	8.
10	100	95	37.28	121	20	NaN	NaN	NaN	NaN	22	NaN	NaN	NaN	NaN	NaN	32	NaN	7.
11	95	100	NaN	89	62.33	NaN	18	NaN	NaN	22	NaN	NaN	NaN	NaN	8	19	70	7.
12	86	96	38	111	66	49	17	NaN	1	27	NaN	7.39	45	95	NaN	16	NaN	Ni
13	88	100	36.3	99	66	52	16	NaN	-3	20	1	7.35	39	NaN	NaN	14	NaN	Ni
14	116	97	38.28	200	108	90	24	NaN	6	NaN	0.7	7.51	39	NaN	NaN	NaN	NaN	Ni
15	110	99	36.4	116	219	66	19	NaN	-8	19	NaN	7.22	46	96	NaN	85	NaN	Ni
16	54	95	NaN	103	63	NaN	11	NaN	NaN	30	NaN	NaN	NaN	NaN	NaN	11	NaN	9.
17	98	94	NaN	95	62	45	15	NaN	NaN	26	NaN	NaN	NaN	NaN	12	11	55	7.
18	72	96	NaN	103	62	45	20	NaN	-1	NaN	NaN	7.4	36	98	NaN	NaN	NaN	Ni
19	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	24	NaN	NaN	NaN	NaN	NaN	65	NaN	9.
20	84	98	NaN	106	67.33	NaN	29	NaN	NaN	20	NaN	NaN	NaN	NaN	NaN	37	NaN	6.
21	69.5	100	37.17	109.5	71	NaN	17.5	NaN	NaN	24	NaN	NaN	NaN	NaN	NaN	14	NaN	Ni

Messages: 4 User: stefaniegreay



When you click the “Code/Results” or “Split” tab and then “Results,” you can see the contents of the dataset, to verify the number of observations and variables are as expected.

The screenshot shows the SAS Studio interface. On the left, the 'Server Files and Folders' pane lists various files, with 'dataSepsis.csv' selected. The main window displays the 'Results' tab for the 'dataSepsis' dataset. The 'Table of Contents' section shows 'The CONTENTS Procedure' output, which includes a table of dataset properties and a table of engine/host dependent information.

Property	Value
Data Set Name	WORK.IMPORT
Member Type	DATA
Engine	V9
Created	07/10/2021 13:19:25
Last Modified	07/10/2021 13:19:25
Protection	
Data Set Type	
Label	
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRU64, LINUX_IA64
Encoding	utf-8 Unicode (UTF-8)
Observations	36302
Variables	41
Indexes	0
Observation Length	168
Deleted Observations	0
Compressed	NO
Sorted	NO

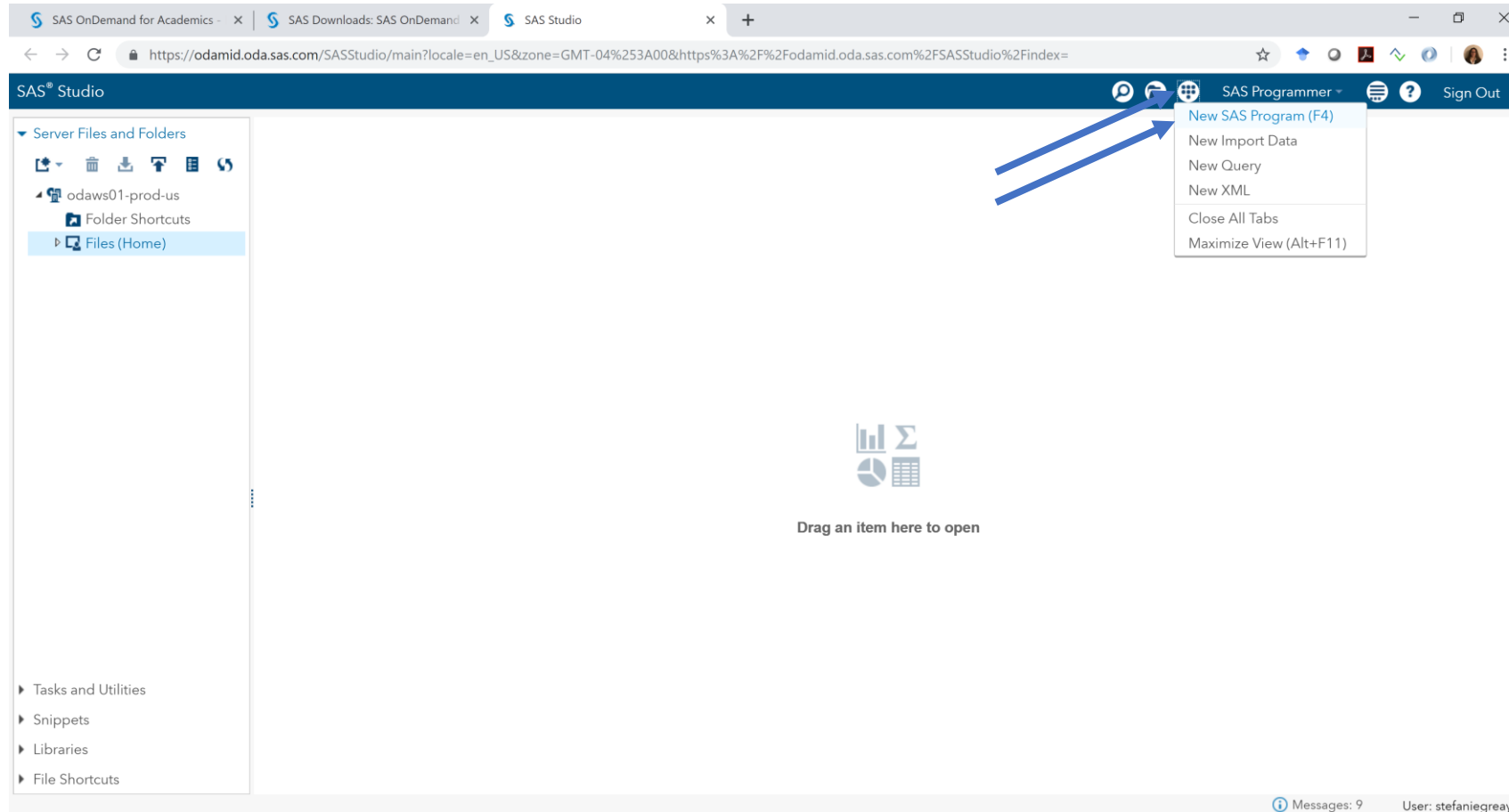
Property	Value
Data Set Page Size	131072
Number of Data Set Pages	47
First Data Page	1
Max Obs per Page	779
Obs in First Data Page	736
Number of Data Set Repairs	0
Filename	/sas/work/SAS_work6E5F0000DA42_odswe03-usw2.ods.sas.com/SAS_work75820000DA42_odswe03-usw2.ods.sas.com/Import.sas7bdat
Release Created	9/04/2016
Host Created	Linux
Inode Number	537069780
Access Permission	rw-r--r--
Owner Name	stefaniegreay
File Size	6MB
File Size (bytes)	6291456

#	Variable	Type	Len	Format	Informat
15	AST	Char	3	\$3.	\$3.
35	Age	Num	8	BEST12.	BEST32.
17	Alkalinephos	Char	3	\$3.	\$3.

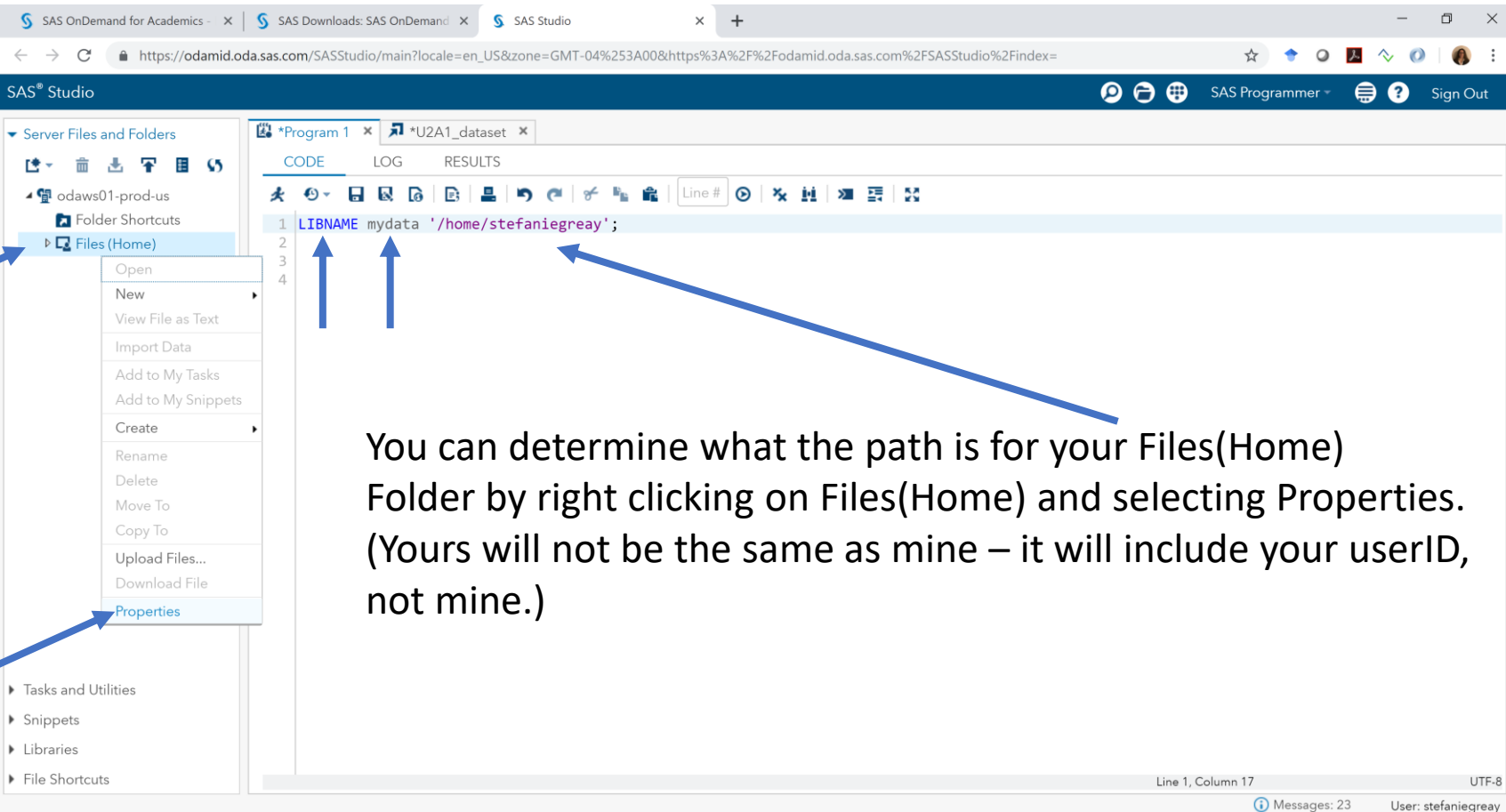
At the bottom right, the status bar shows 'Messages: 4' and 'User: stefaniegreay'.



To get started working with the dataset you just imported, start a new SAS program.



To create a SAS Library for your Files(Home) folder, you need to use a libname statement



The screenshot shows the SAS Studio web interface. On the left, the 'Server Files and Folders' pane displays a tree structure with 'Files (Home)' selected. A right-click context menu is open over 'Files (Home)', with the 'Properties' option highlighted at the bottom. A blue arrow points from the 'Properties' option to the text box. The main editor area shows a SAS program with the following code:

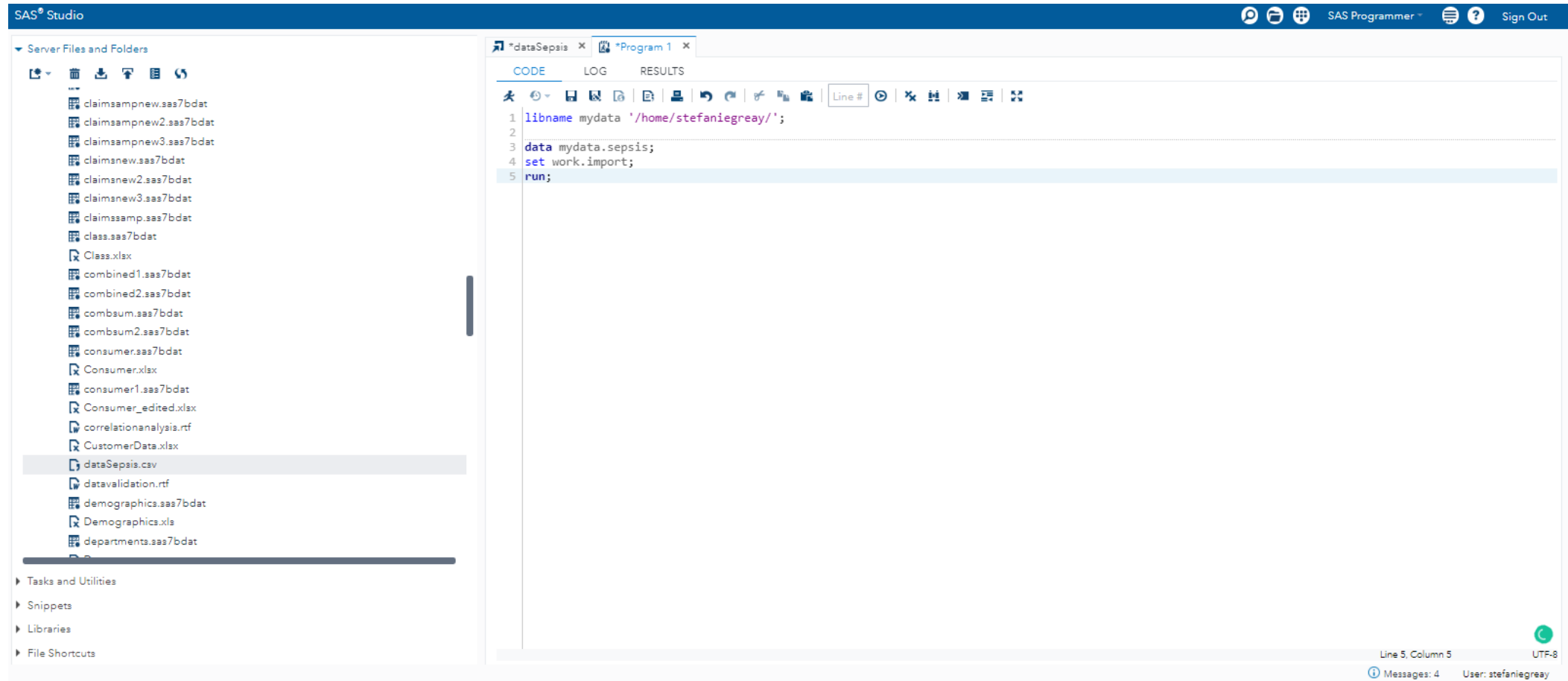
```
1 LIBNAME mydata '/home/stefaniegreay';  
2  
3  
4
```

Two blue arrows point from the text box to the words 'LIBNAME' and 'mydata' in the code. A third blue arrow points from the text box to the path '/home/stefaniegreay'.

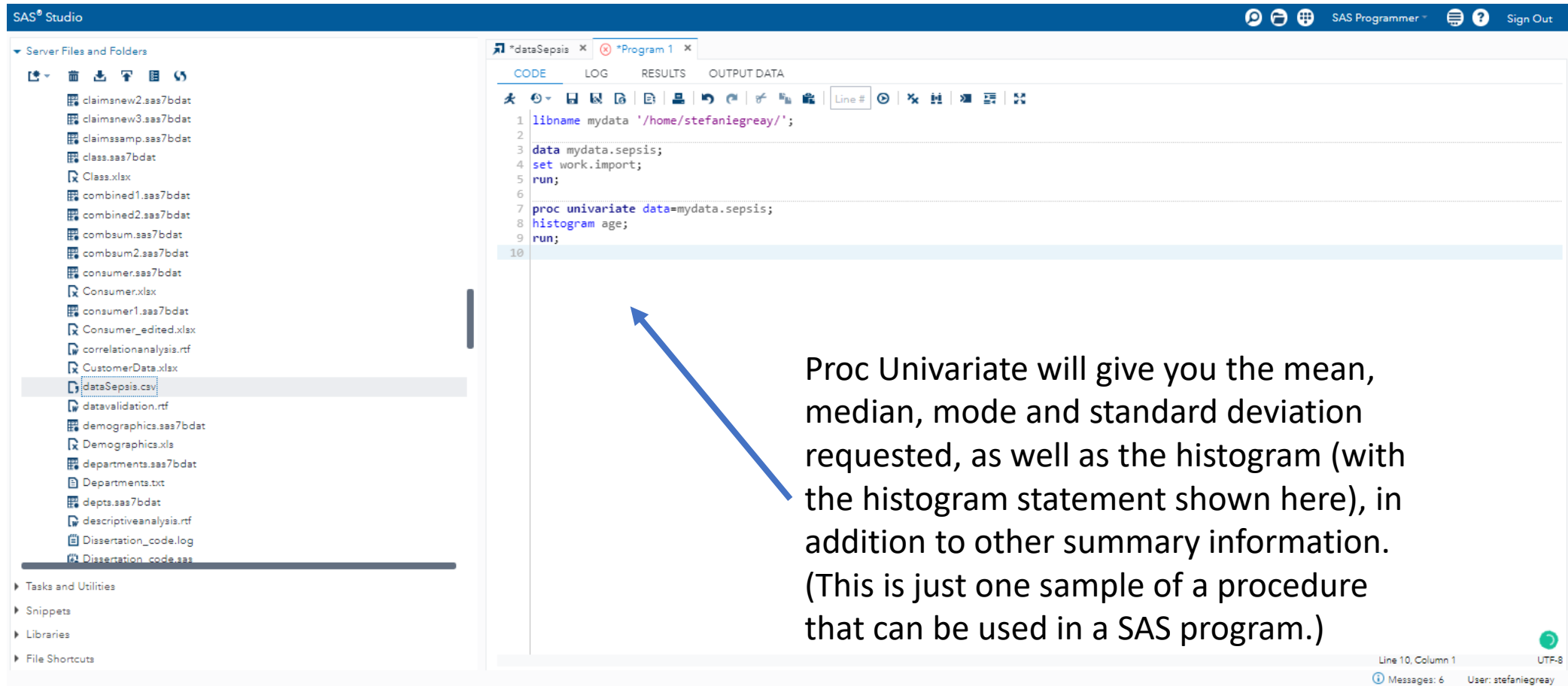
You can determine what the path is for your Files(Home) Folder by right clicking on Files(Home) and selecting Properties. (Yours will not be the same as mine – it will include your userID, not mine.)



Save the temporary SAS dataset created by the import to your library using the following sample code.



You can now run any procedures against that dataset via the code window.



The screenshot displays the SAS Studio interface. On the left, the 'Server Files and Folders' pane lists various files, with 'dataSepsis.csv' highlighted. On the right, the 'CODE' window shows a SAS program with the following lines:

```
1 libname mydata '/home/stefaniegreay/';  
2  
3 data mydata.sepsis;  
4 set work.import;  
5 run;  
6  
7 proc univariate data=mydata.sepsis;  
8 histogram age;  
9 run;  
10
```

A blue arrow points from the text box to the code editor. The status bar at the bottom right indicates 'Line 10, Column 1', 'UTF-8', 'Messages: 6', and 'User: stefaniegreay'.

Proc Univariate will give you the mean, median, mode and standard deviation requested, as well as the histogram (with the histogram statement shown here), in addition to other summary information. (This is just one sample of a procedure that can be used in a SAS program.)



Code template

```
libname mydata '/home/userid/';
```

```
data mydata.sepsis;
```

```
set work.import;
```

```
run;
```

```
proc univariate data=mydata.sepsis;
```

```
histogram age;
```

```
run;
```

